

Data Management at NIST

John Henry J. Scott¹

Raymond Plante

Office of Data and Informatics (ODI)

Material Measurement Laboratory

National Institute of Standards and Technology

Tuesday, March 20, 2018

Scientific Data Management for Photon and Neutron Facilities

HZB, Berlin

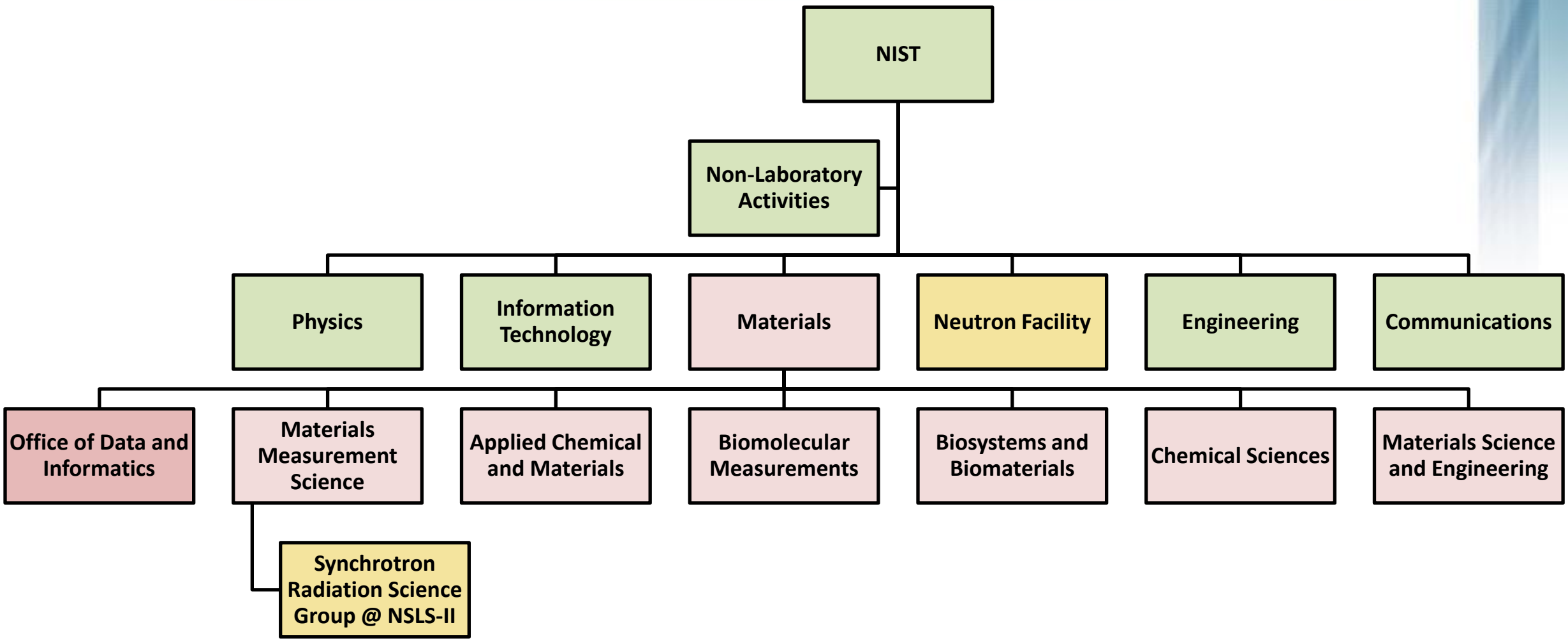
Credits

- ODI Staff and Details: Gretchen Greene, Adam Morey, Chandler Becker, Peter Linstrom, Arlin Stoltzfus, Kim Tryka, Andrea Medina-Smith, Zachary Trautt, June Lau, et al.
- ITL Staff: Mary Brady, Alden Dima, Sharief Youssef, et al.
- Others: Office of Information Systems Mgt, Kathy Sharpless, many others

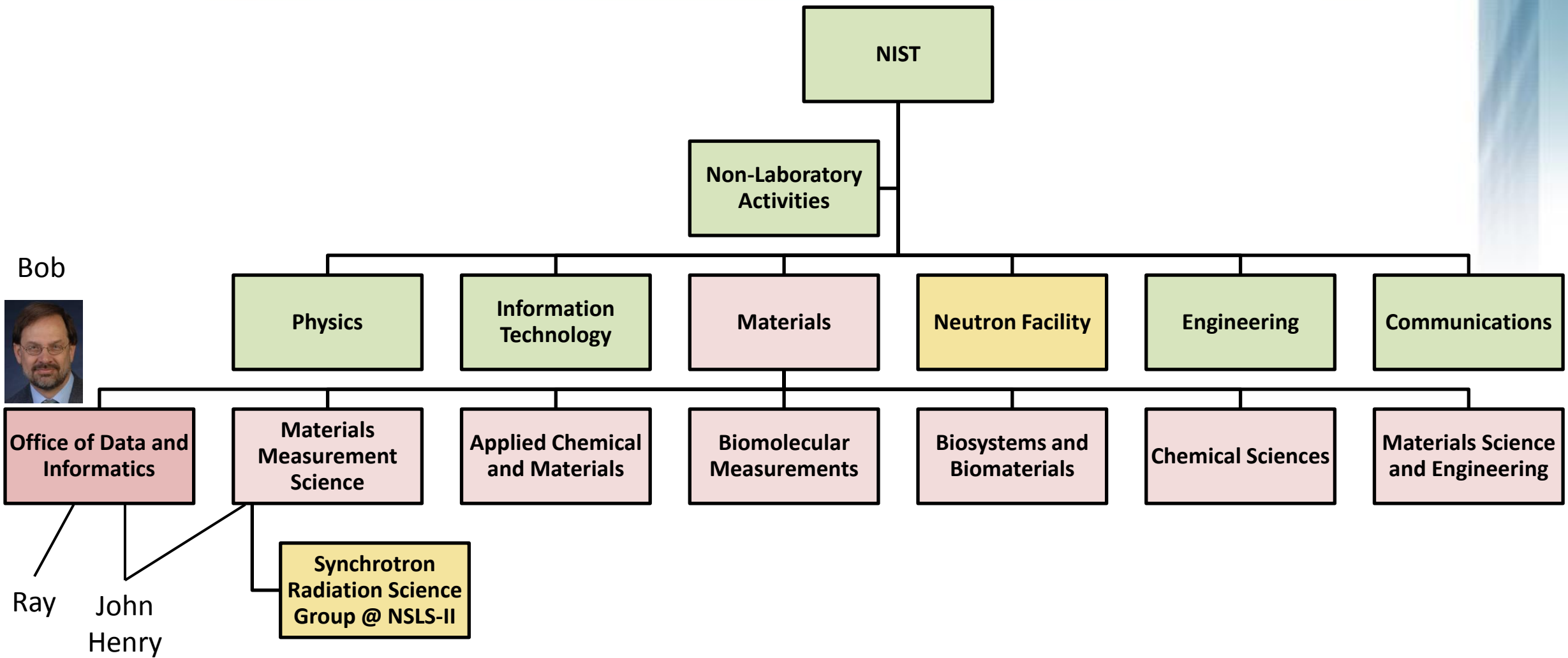
Challenges

- Incredible diversity of research programs
 - Size, complexity, discipline
 - Most data generated by small research groups (long tail)
- Conservative culture, resistant to change
 - Must demonstrate high ratio of benefit to cost
- Little recognition for data sharing/publication for advancement
- Proprietary data formats
 - Many bench instruments have opaque formats
 - Vendors want to sell proprietary software as well as the instrument
- IT security
 - Widely-used tools like Dropbox prohibited
 - New deployments require A&A, ATO – takes time and effort

NIST Organization



NIST Organization



**MATERIAL
MEASUREMENT
LABORATORY**

Office of Data and Informatics

Standard Reference Data

- Distribution
- Sales
- Infrastructure
- Usage analysis and impact
- Improve web sites and user interfaces
- Provide APIs

Research Data

- **Improve data management practices**
- Data management planning tools
- **Laboratory automation**
- Electronic Lab Notebooks
- **NIST open data repository**
- **NIST data portal**

Data Science

- Informatics and analytics resource
- Liaison with NIST Information Technology Laboratory
- Big data
- Cloud computing
- National Strategic Computing Initiative

Community

- National Data Services Consortium
- Research Data Alliance
- Other US federal agencies (NIH, DOE, NSF)
- CENDI
- National Metrology Institutes (NMIs)
- BIPM (Paris)
- CODATA, WDS



Robert Hanisch
Director, Office of Data and Informatics

Data Management at NIST

Discover



- Standard Reference Data
- Materials Data Repository
- Materials Data Facility
- Persistent identifiers (DOIs, handles)

- Materials Resource Registry (data, code)
- International Metrology Resource Registry
- NIST Enterprise Data Inventory
- data.gov
- NIST Public Data Repository and Search Portal



Access

Interoperate

- Materials Data Curator
- Data type registry
- Schema repository
- Lab info mgmt systems



Discovery

Materials Resource Registry

NIST Home Services > Login Help Contact

Materials Resource Registry

Part of the Materials Genome Initiative

[SEARCH FOR RESOURCES](#) [ADD YOUR RESOURCE](#)

Find Materials Data

This system allows for the registration of materials resources, bridging the gap between existing resources and the end users. The Materials Resource Registry functions as a centrally located service, making the registered information available for research to the materials community.

This is being developed at the National Institute of Standards and Technology and is made available to solicit comments from the Material Science community. Please do not enter any proprietary data into this system.

Home Page

- Services
- Search for resources
- Add your resource
- Login
- Help
- Contact

<https://materials.registry.nist.gov/>

<http://imrr.bipm.org/>



International Metrology Resource Registry



SEARCH FOR RESOURCES

ADD YOUR RESOURCE

Find Resources

This system allows for the registration of resources, bridging the gap between existing resources and the end users. The International Metrology Resource Registry functions as a centrally located service, making the registered information available for research to the global community.

This is being developed at the Bureau International des Poids et Mesures and is made available to solicit comments from the global community. Please do not enter any proprietary data into this system.

Home Page

Services

Search for resources

Add your resource

Dashboard

Logout

Help

Contact

Data Discovery for Public Research Data

The screenshot displays the NIST Science Data Portal interface. At the top left is the NIST logo and the text "Science Data Portal" with a "1.0.0-beta" badge. A navigation bar contains links for "Key Datasets", "Standard Reference Data (SRDs)", "Developer", "About", and "Find Papers". The main header features the title "NIST Data Discovery" and the tagline "Explore data, tools, and resources for Science, Engineering, Technology and more". Below this is a search bar with the query "SRD 101", a dropdown menu set to "All Research", a "Search" button, and a link to "Advanced Search". Example search terms are listed below the search bar. The "FEATURED DATA DOMAINS" section contains eight blue buttons for: INFORMATION TECHNOLOGY, MATHEMATICS AND STATISTICS, MANUFACTURING, FORENSICS, MATERIALS, PHYSICS AND NEUTRON, ADVANCED COMMUNICATIONS, and CHEMISTRY. The footer includes the NIST logo, full name, address, and social media icons for Twitter, Facebook, Google+, YouTube, RSS, and Email.

NIST Science Data Portal

1.0.0-beta

Home Key Datasets Standard Reference Data (SRDs) Developer About Find Papers

NIST Data Discovery

Explore data, tools, and resources for Science, Engineering, Technology and more

Q "SRD 101" All Research Search Advanced Search

Examples: ["Kinetics database"](#) [Gallium](#) ["SRD 101"](#) [XPDB](#) [Interatomic Potentials](#)

FEATURED DATA DOMAINS

- INFORMATION TECHNOLOGY
- MATHEMATICS AND STATISTICS
- MANUFACTURING
- FORENSICS
- MATERIALS
- PHYSICS AND NEUTRON
- ADVANCED COMMUNICATIONS
- CHEMISTRY

NIST National Institute of Standards and Technology
U.S. Department of Commerce
HEADQUARTERS
100 Bureau Drive
Gaithersburg, MD 20899



Data Discovery for Public Research Data

NIST Science Data Portal 1.0.0-beta

Key Datasets ▾ Standard Reference Data (SRDs) Developer ▾ About ▾ Find Papers ▾

NIST Data Discovery

Explore data, tools, and resources for Science, Engineering, Technology and more

🔍 "SRD 101" All Research ▾ Search [Advanced Search](#)

Examples: ["Kinetics database"](#) [Gallium](#) ["SRD 101"](#) [XPDB](#) [Interatomic Potentials](#)

Search NIST public data records

- View metadata
- Filter results
- Access data files, metadata
- APIs allow interoperability with client tools
- Records link to Public Data Repository

MATHEMATICS AND STATISTICS MANUFACTURING FORENSICS

PHYSICS AND NEUTRON ADVANCED COMMUNICATIONS CHEMISTRY

HEADQUARTERS
100 Bureau Drive
Gaithersburg, MD 20899

🐦 f G+ YouTube 📡 ✉

PHYSICAL MEASUREMENT LABORATORY

About PML +

Divisions +

Products/Services +

News/Multimedia

Programs/Projects

Facilities +

Email Newsletter

Atomic Spectra Database



Version 5

[Version History & Citation Information](#) | [Disclaimer](#)

Welcome to the NIST Atomic Spectra Database, NIST Standard Reference Database #78. The spectroscopic data may be selected and displayed according to wavelengths or energy levels by choosing one of the following options:

LINES

Spectral lines and associated energy levels displayed in wavelength order with all selected spectra intermixed or in multiplet order. Transition probabilities for the lines are also displayed where available.

LEVELS

Energy levels of a particular atom or ion displayed in order of energy above the ground state.

**GROUND STATES &
IONIZATION ENERGIES**

Ground states and ionization energies of atoms and atomic ions.

Data Discovery for Public Research Data

NIST Science Data Portal

1.0.0-beta

[Home](#)
[Key Datasets](#)
[Standard Reference Data \(SRDs\)](#)
[Developer](#)
[About](#)
[Find Papers](#)

All Research
Search
[Advanced Search](#)

Examples: ["Kinetics database"](#) [Gallium](#) ["SRD 101"](#) [XPDB](#) [Interatomic Potentials](#)

Filters Clear All

- Resource Type
 - Public Data Resource **4**
 - SRD **1**
- Research Topics
 - Neutron Research **2**
 - Standards **1**
 - Unspecified **1**
- Record has
 - Data File **4**

Authors and Contributors

Keyword

4 records found Customize Fields

|< < **1** > >|

NIST Center for Neutron Research raw data archive
 Neutron scattering data from NCNR's thermal and cold neutron scattering instruments.
Subject Keywords: *neutron research, neutron scattering, neutron diffraction, neutron reflectometry, small-angle neutron scatteri...* [Read more](#)
[Visit Home Page](#)

Source files for online unpolarized neutron and x-ray reflectivity calculator
 The data stored here is a program to be run in a web browser, and as such the end-user must have access to the javascript code.
Subject Keywords: *neutron reflectometry reflectivity calculator web javascript d3*
[Visit Home Page](#)

FIZ/NIST Inorganic Crystal Structure Database (ICSD) - SRD 84
 The Inorganic Crystal Structure Database (ICSD) is produced cooperatively by the Fachinformationszentrum Karlsruhe (FIZ) and the National Institute of Standards and Technology (NIST). Components and devices used in a broad spectrum of technology sectors such as health c... [Read more](#)
Subject Keywords: *Rietveld profiles, X ray crystallography, X ray diffraction, X rays, XRD, absolute configurations, chemical st...* [Read more](#)
[Visit Home Page](#)

test without CR
 Code is developed in python, html, css, xml and javascript and shared via GitHub's USNISTGOV organization.
Subject Keywords: *GitHub pages template*

NIST Public Data Repository – Basic Landing Page

Public Data Resource

FIB SEM image data set of *Caenorhabditis elegans* exposed to 60 nm Au nanoparticles

Contact: [Keana Scott](#) .. 

Identifier: [doi:10.18434/M3C09F](https://doi.org/10.18434/M3C09F)

Last modified: **2015-09-20**

Description

This folder contains image data sets from 14 separate serial sectioning sessions. The entire data folder consists of 1379 8 bit tif images and is 47.3 GB in size. Serial sectioning was performed using FEI Helios 660 NanoLab focused ion beam scanning electron microscope (FIB SEM) and Auto Slice and View G3 software. The sample was a heavy metal stained and resin embedded *Caenorhabditis elegans* (*C. elegans*) that were exposed to 60 nm Au nanoparticles. Detailed descriptions of the worm preparation and resin block processing are described in Johnson, M.E. et al. (*ACS Nano*, 2016). Although the images were collected over 14 different sessions, they represent a contiguous section of a worm.


Subject Keywords: Advanced Materials, Biosciences and Health, Environment and Climate, Manufacturing, Visualization Research, Nanotechnology

References:




This data is referenced in :

 <http://pubs.acs.org/doi/abs/10.1021/acsnano.6b06582>


Access To Data:

 This data is public.

Access

-  [Visit Home Page](#)
-  [Download all data](#)
-  [Export JSON](#)

Use

- » [Cite this resource](#)
-  [License Statement](#)

Find



-  [Similar Resources](#)
-  [Resources by Authors](#)

Table of Contents

- [Description](#)
- [Files](#)
- [Metadata](#)

NIST Public Data Repository – Basic Landing Page

Public Data Resource

FIB SEM image data set of *Caenorhabditis elegans* exposed to 60 nm Au nanoparticles

Contact: [Keana Scott](#) ..

Identifier: [doi:10.18434/M3C09F](https://doi.org/10.18434/M3C09F)

Last modified: 2015-09-20

Description

This folder contains image data sets from 14 separate serial sectioning sessions. The entire data folder consists of 1379 8 bit tif images and is 47.3 GB in size. Serial sectioning was performed using FEI Helios 660 NanoLab focused ion beam scanning electron microscope (FIB SEM) and Auto Slice and View G3 software. The sample was a heavy metal stained and resin embedded *Caenorhabditis elegans* (*C. elegans*) that were exposed to 60 nm Au nanoparticles. Detailed descriptions of the worm preparation and resin block processing are described in Johnson, M.E. et al. (ACS Nano, 2016). Although the images were collected over 14 different sessions, they represent a contiguous section of a worm.


Subject Keywords: Advanced Materials, Biosciences and Health, Environment and Climate, Manufacturing, Visualization Research, Nanotechnology

References:

This data is referenced in :

<http://pubs.acs.org/doi/abs/10.1021/acsnano.6b06582>

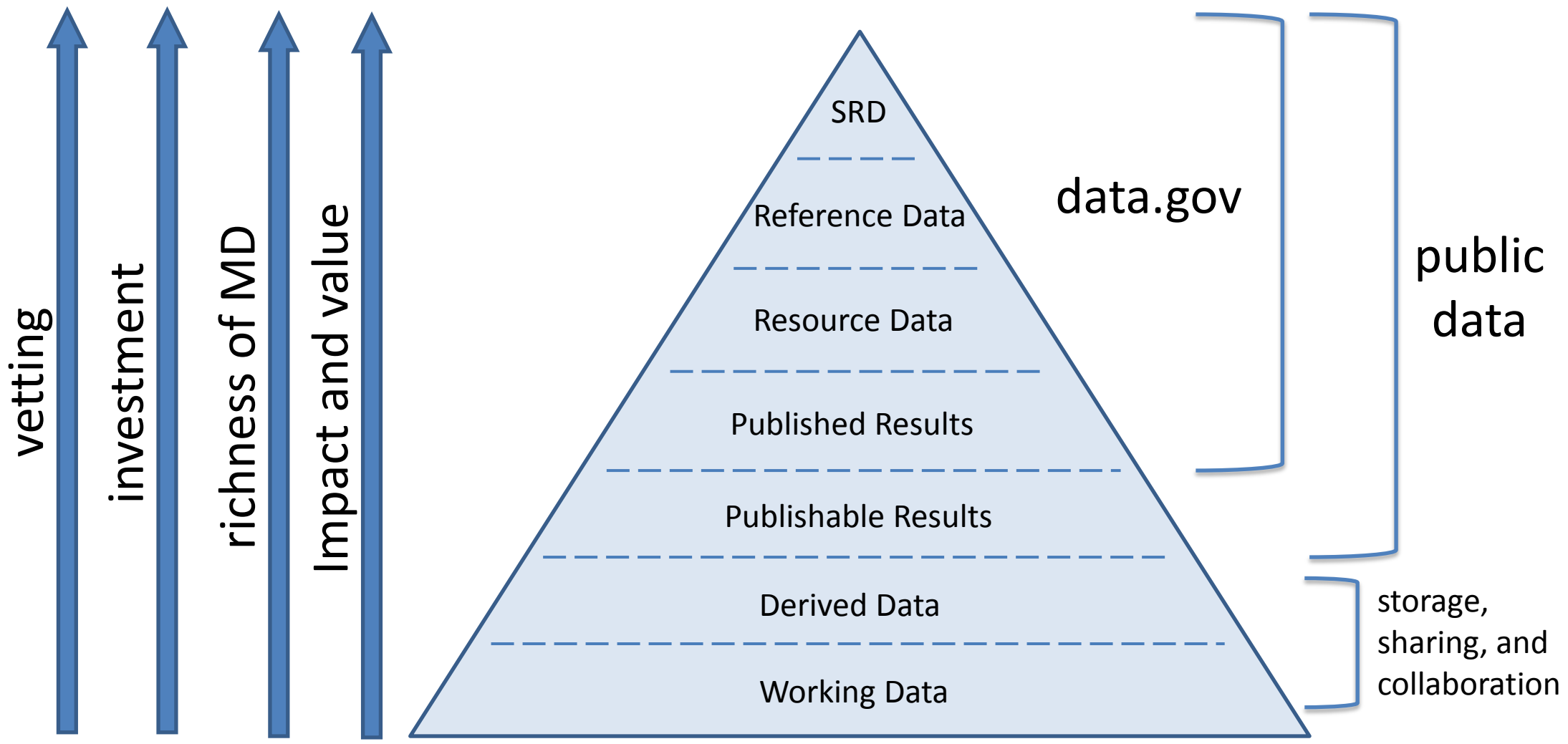
Access To Data:

 This data is public.

- Published data landing page example
- The Public Data Repository provides capabilities for rich Data Publications
- Supports **NIST Extended metadata model** for domain science
- Standards for data publication are implemented

Access

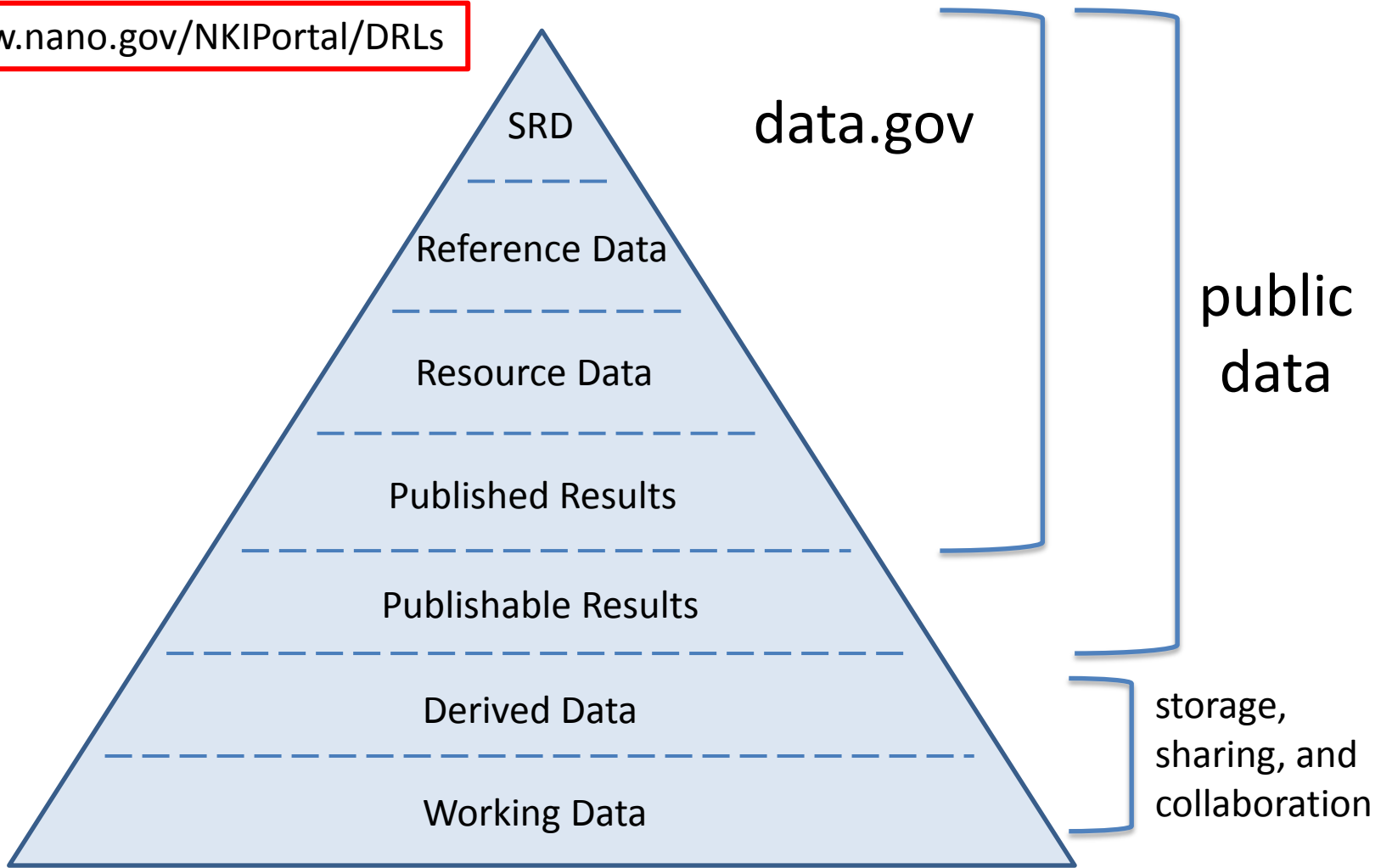
NIST Public Data Access Policy



NIST Public Data Access Policy

Data Readiness Level (DRL)
0. Invalid data
1. Raw or unscaled data
2. Scaled data
3. Scaled data with defined precision <u>or</u> noise level
4. Scaled data with defined precision <u>and</u> noise levels, but <u>not related</u> to the larger body of scientific knowledge
5. DRL 4 data <u>related</u> to the larger body of scientific knowledge, but with measurement uncertainty too large for data standards
6(X). Standards-quality data of X % measurement uncertainty

<https://www.nano.gov/NKIPortal/DRLs>





Creating a Data Management Plan

Main Menu

- Site
 - FAQ
 - Feedback
 - Resources
- Data Management Plan
 - Create DMP
 - My Org DMPs
 - Browse All DMPs
 - About DMPs
- EDI Dataset
 - Register New Dataset
 - My Org Records (39)
- Reports
 - Metadata Report
 - JSON Export

Welcome to the guided DMP!

Would you like to register a new Data Management Plan (DMP) or edit an existing DMP?

[Guided Registration](#) [Form Registration](#) [Edit Existing DMP](#)

DMP Guided

Basic Information Personnel Keywords Data Description Data Categories Data Preservation Reviewer

Title*:

Project Identifier:

[Save](#) [Next](#)

1. DMP* 2. Title* 3. Contact(s)* **4. Description*** 5. Keywords* 6. Data Use 7. Access 8. Files (Beta)

* Enter a description of your dataset:

Human-readable description (e.g., an abstract) with sufficient detail to enable a user to quickly understand whether the asset is of interest.

* Language: English

* When was the last time this dataset was updated?

* Is this dataset updated on a regular basis? Yes No

Back Save Next

Creating an Enterprise Data Inventory Record

1. DMP* 2. Title* 3. Contact(s)* 4. Description* 5. Keywords* 6. Data Use 7. Access **8. Files (Beta)**

Browse and attach files associated with this dataset.

Attached Files

File Name	File Size
Upload a File: <input type="text"/> Browse...	

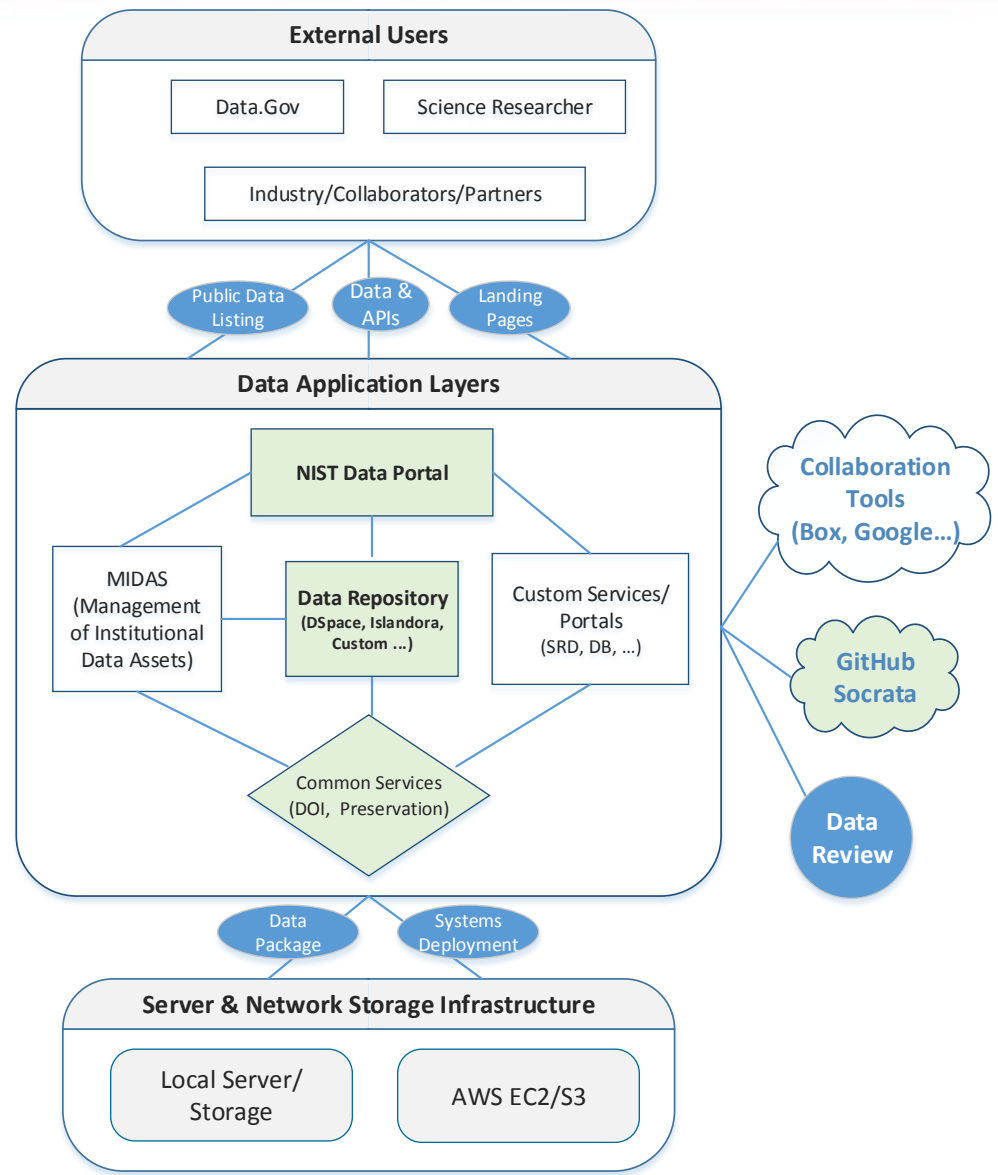
File size cannot exceed 40MB and cannot include binary files (.exe, .dll).

[Attach File](#)

Note: This portion of the MIDAS application is a beta version and has limited functionality and file upload restrictions.

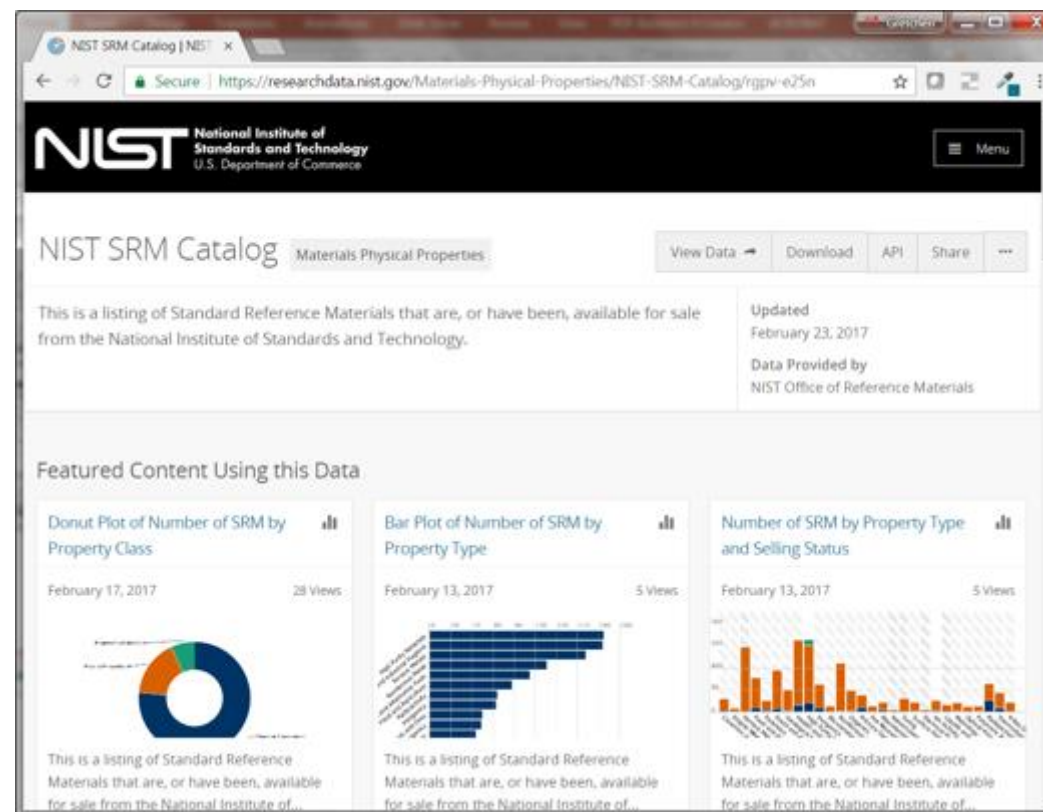
Back Save

Research Data Infrastructure



Sharing with Socrata Publishing Tool

- NIST public data records hosted in external Socrata publishing system
- Provides web landing page view and catalog for tabular datasets
- Search, filtering, and visualization tools
- Auto-generated API allows users to access data by script rather than by website



NIST[NIST Time](#) | [NIST Home](#) | [About NIST](#) | [Contact Us](#)

Material Measurement Laboratory

[About MML](#) ▾[Publications](#)[Topic/Subject Areas](#) ▾[Products/Services](#) ▾[News/Multimedia](#)[Events](#)

Material Measurement Laboratory Repository Server

This is the NIST Material Measurement Laboratory data repository server.

Use of this server is subject to [terms of service](#)

The repository itself is located at <https://materialsdata.nist.gov/dspace/xmlui> and may be accessed using the button below.

To get an account on this system (required for uploading), send a message to the [administrator](#). Please include your requested username, e-mail address, first and last name, the name of the community you wish to access (if known), and if possible a brief explanation for your reasons for requesting an account.

If you are new to using repositories or DSpace in specific, you may wish to read either the [short](#) or [in-depth](#) FAQs. Links to these have also been provided below.

[Request an Account \(for Uploading\)](#)

[View the Repository](#)

[Manage Your Credentials](#)

NIST

Material Measurement Laboratory

materialsdata.nist.gov

NIST Repositories

Communities in NIST Repositories

Select a community to browse its collections. [R] indicates an invitational community, [Z] indicates an archived community.

- [ASM Structural Materials Data Demonstration Project](#)
- [Chemical Sciences](#)
- [CHiMaD Data Collections \[R\]](#)
- [Community for Greenhouse Gases](#)
- [Computational File Repository](#)
- [Experimental Data Repository](#)
- [Genome in a Bottle](#)
- [Heusler Phases: First Principles Simulations \[R\]](#)
- [ICME Approach to Development of Lightweight 3GAHSS Vehicle Assembly \[R\]](#)
- [ICME of Carbon Fiber Composites for Lightweight Vehicles \[R\]](#)
- [MGI Catalogs](#)
- [MICCoM Collections \[R\]](#)
- [NanoRelease \[R\]](#)
- [NIST/DOE-EERE Advanced Automotive Cast Magnesium Alloys \[R\]](#)
- [NIST Thermodynamics and Kinetics Test Space \[R\]](#)
- [Porous Metals and Ceramics: Freeze-casting under microgravity and terrestrial conditions](#)
- [RDA Demonstration Project: DTR/PID & MGI Infrastructure \[R\]](#)
- [RVE fracture VUMAT for QP980 steel](#)
- [State Variable Model for QP980](#)
- [Synchrotron Studies of Slot Die Coated Films \[R\]](#)
- [Thermal Conductivity of CVD Diamond - DARPA Round Robin \[R\]](#)
- [TMS Springer Integrating Materials and Manufacturing Innovation \(IMMI\)](#)

Search NIST Repositories

[Advanced Search](#)

Browse

All of NIST Repositories
[Communities & Collections](#)
[By Issue Date](#)
[Authors](#)
[Titles](#)
[Subjects](#)

My Account

[Login](#)

Discover

Author

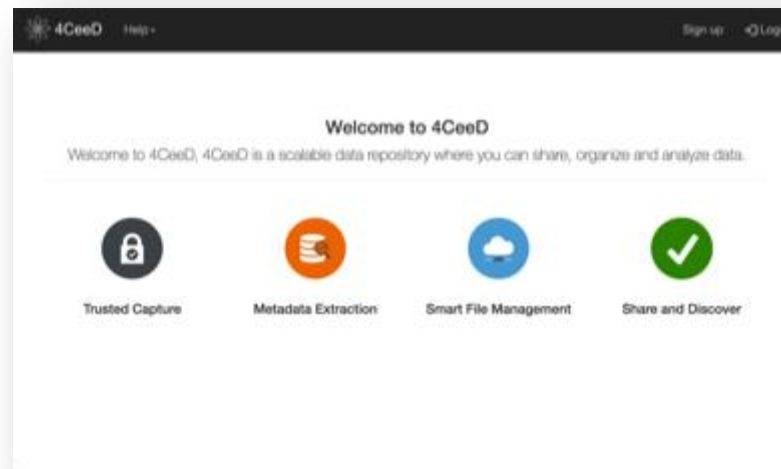
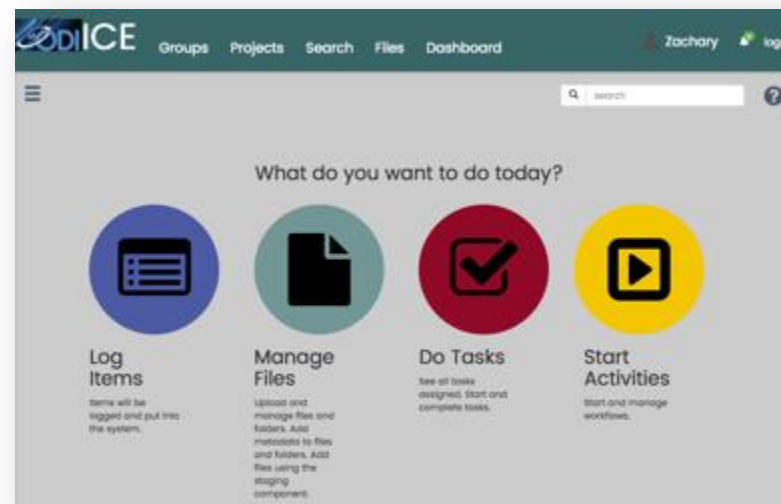
[Du, Y. \(10\)](#)
[Abu-Farha, Fadi \(8\)](#)
[Bower, Allan F. \(8\)](#)
[Burton, Benjamin P. \(7\)](#)
[van de Walle, Axel \(7\)](#)
[Xu, Honghui \(7\)](#)
[Zhang, L. \(7\)](#)
[Liu, Shuhong \(6\)](#)
[DeSchepper, Daniel C. \(5\)](#)
[Flanagan, David P. \(5\)](#)
[... View More](#)

Subject

Interoperability

Laboratory Information Management Systems

- Integrated Collaborative Environment (ICE), a.k.a. Hyperthought
 - Running now at <http://ice.nist.gov>
 - Developed by Air Force Research Laboratory
- Timely and Trustworthy Curating and Coordinating Data Framework (T2C2) 4CeeD system
 - Running now at <http://t2c2.nist.gov:32500/>
 - Developed by University of Illinois at Urbana-Champaign
- ~~• Also considering Discovery Environment for Relational Information and Versioned Assets (DERIVA) from USC~~



Laboratory Information Management Systems

- Capture instrument metadata at the source
 - Metadata extractors
 - Often must reverse engineer proprietary binary formats
- Move experiment metadata into database
 - Enable search across many experiments
 - Do not use filenames/file system for metadata storage
- Enable scripted data processing, calibration, feature extraction
- Support data management from acquisition to publication; improve reproducibility

Materials Data Curation System

Materials Data Curation System

Part of the Materials Genome Initiative

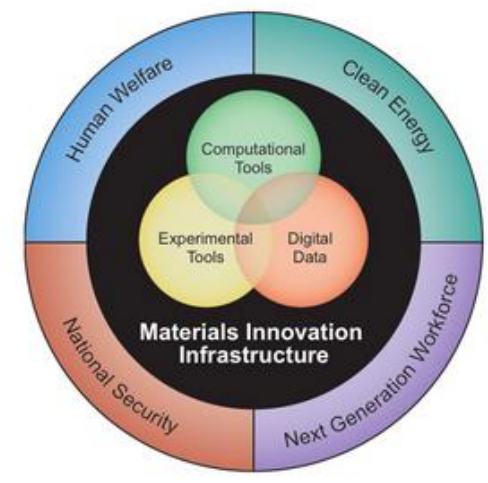
[Login](#) | [My Profile](#) | [Help](#)

[Home](#)

Materials Data Curator

This system allows for the curation of Material Data in a repository using predefined templates.

This is being developed at the National Institute of Standards and Technology and is made available to solicit comments from the Material Science community. Please do not enter any proprietary data into this system.



Available Options

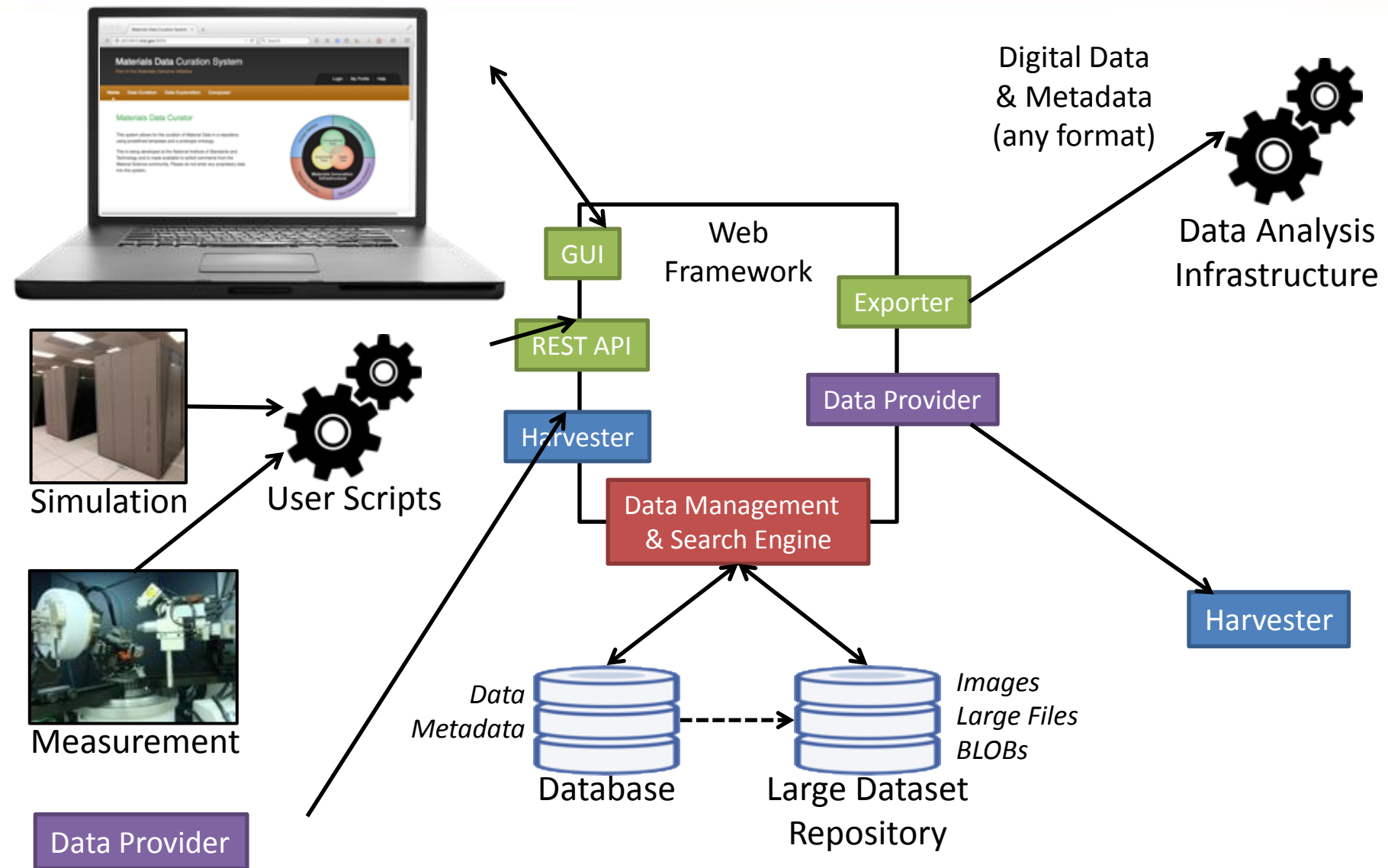
[All Options »](#)

Most Recent Templates

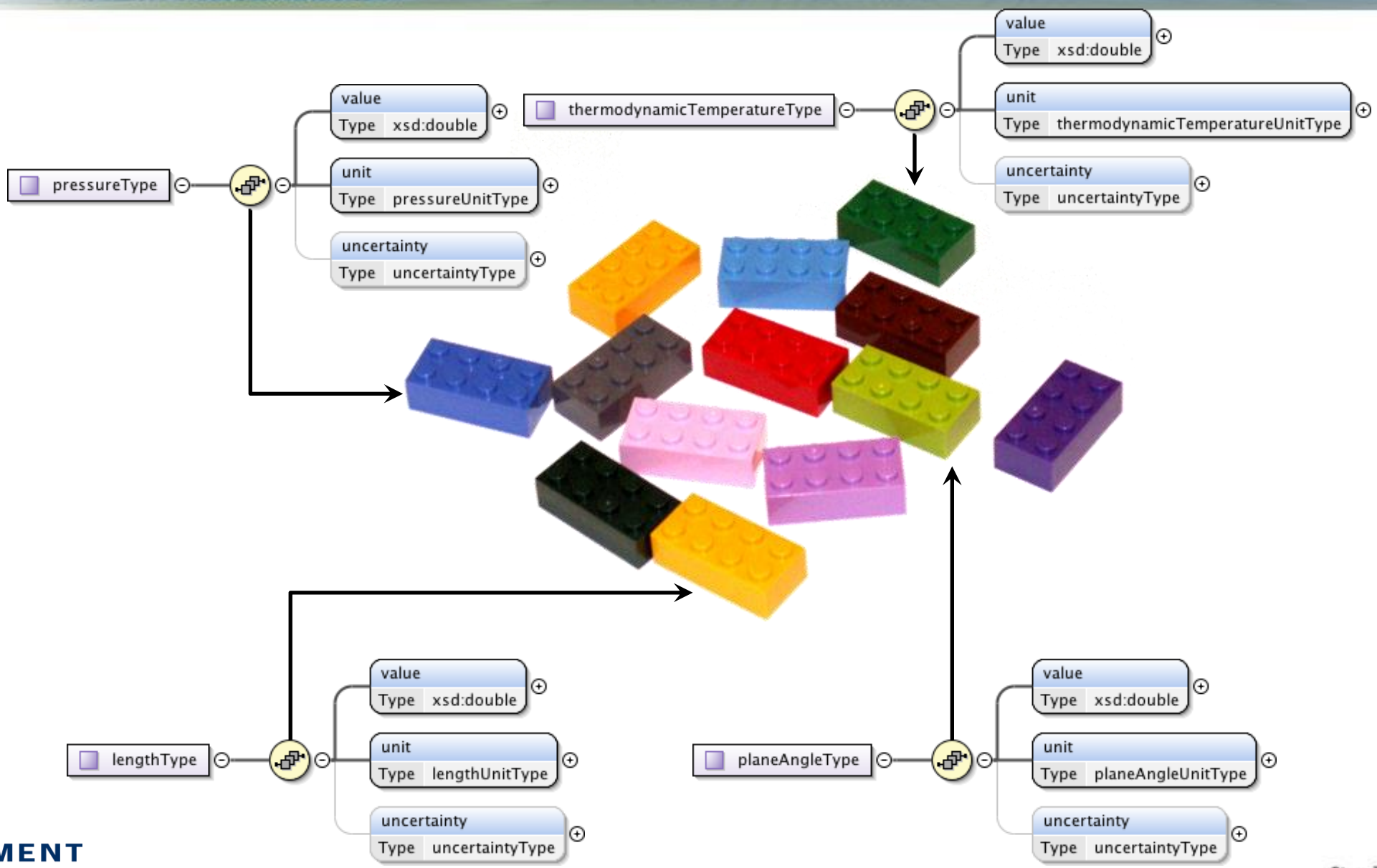
[Browse All »](#)

- species-qs | species.xsd.xml
- TEM-Tutorial-Tamu | workshop-TEM.xsd
- TEM-Tut | workshop-TEM.xsd

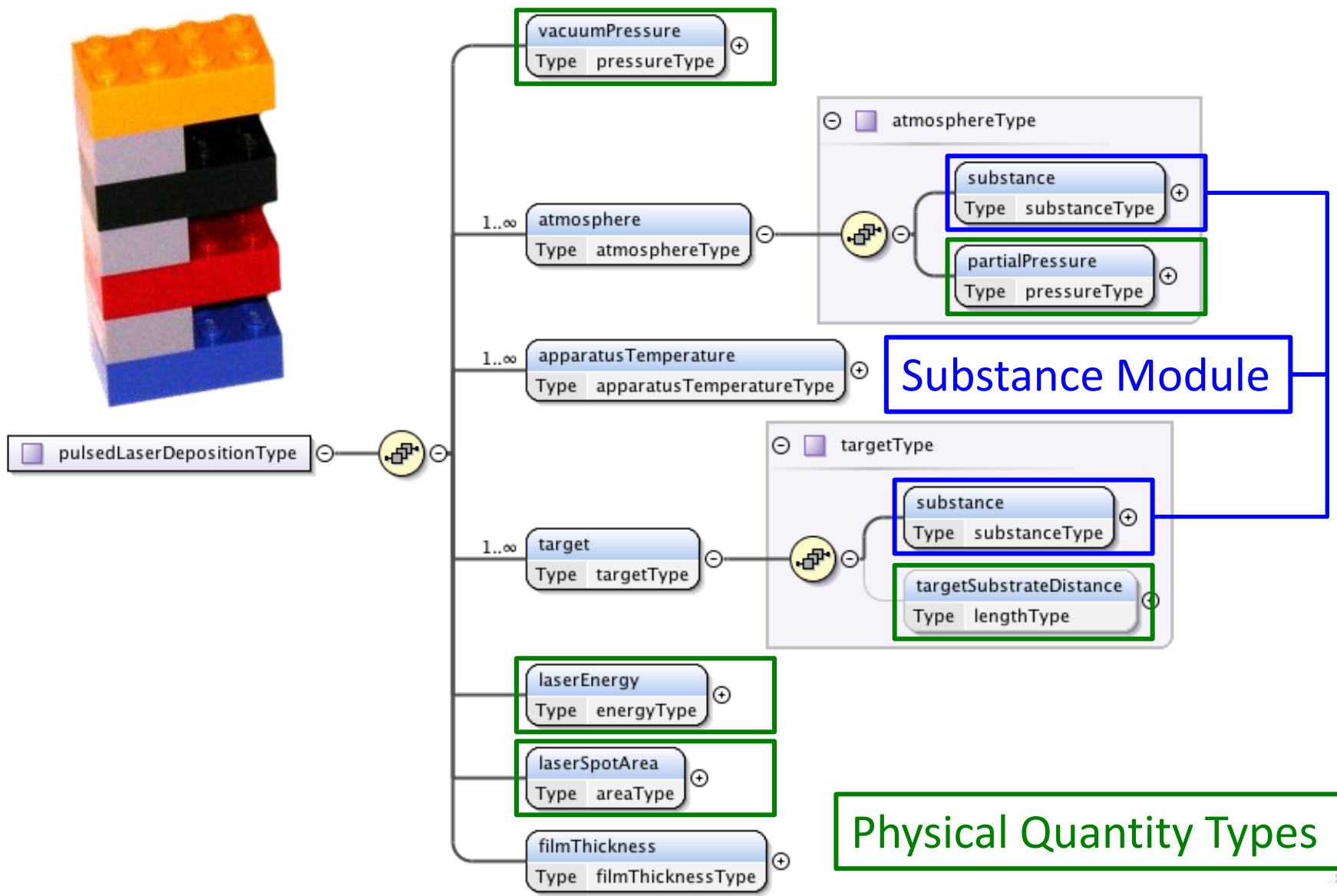
Materials Data Curation System



Modularity: Foundational Types



Data Models Re-Use Components



...and now over to Ray

NIST Public Data Repository

<https://data.nist.gov/>



Public Data Resource

Multiple Encounter Dataset (MEDS-I) - NIST Special Database 32

Contact: [Patricia Flanagan](#) .
Identifier: [ark:/88434/mds0000fbk](#)
Last modified: 2011-07-11

Description

Multiple Encounter Dataset (MEDS-I) is a test corpus organized from an extract of submissions of deceased persons with prior multiple encounters. MEDS is provided to assist the FBI and partner organizations refine tools, techniques, and procedures for face recognition as it supports Next Generation Identification (NGI), forensic comparison, training, and analysis, and face image conformance and inter-agency exchange standards. The MITRE Corporation (MITRE) prepared MEDS in the FBI Data Analysis Support Laboratory (DASL) with support from the FBI Biometric Center of Excellence.

Research Topics: Information Technology; Biometrics
Subject Keywords: face, biometrics, forensic

References:

This data is referenced in :
<https://www.nist.gov/publications/multiple-encounter-dataset-i-meds-i>

Access To Data:

This data is public.

Files

- [NIST_SD32_MEDS-I_face.zip](#)
- [NIST_SD32_MEDS-I.html.zip](#)

Access

- [Visit Home Page](#)
- [Download all data](#)
- [Export JSON](#)

Use

- [Cite this resource](#)
- [License Statement](#)

Find

- [Similar Resources](#)
- [Resources by Authors](#)

Table of Contents

- [Description](#)
- [Files](#)
- [Metadata](#)

Materials Resource Registry

<https://materials.registry.nist.gov/>

Search for Resources

The screenshot shows the Materials Resource Registry search interface. At the top, there is a search bar with the terms 'electrical' and 'properties' entered. Below the search bar, there are several icons representing different resource types: All Resources, Organizations, Data Collections, Datasets, Services, Informational Sites, and Software. The search results are displayed in a list format, with filters for 'TYPE' and 'MATERIAL TYPE'. The results include:

- MatWeb**: <http://www.matweb.com/index.aspx>
Subject keywords: engineering, material properties, database, supplier
"The heart of MatWeb's services is our searchable online database of engineering materials. We have over 115,000 data sheets in our collection and have many powerful search tools available to help our users find the materials information that they need. While we have a variety of services that we offer to companies in the engineering community, our... show more"
- Microelectronics Packaging Materials Database**
CHDAS LLC
<https://chdasdata.com/products/mgmd/>
Subject keywords: electronics packaging materials, material properties
"The MPMD contains data and information on thermal, mechanical, electrical and physical properties of electronics packaging materials, and it is available in a Web-based format. The database is continually updated and expanded. The MPMD contains over 1,025 materials, 388 properties, and contains approximately 22,500 data curves."
- JARVIS - Joint Automated Repository for Various Integrated Simulations**
Kamal Choudhary/National Institute of Standards and Technology, Faical Y. Congo/National Institute of Standards and Technology, Chandler A. Becker, Francesca M. Tavazza/National Institute of Standards and Technology - NIST
<http://www.ctcms.nist.gov/~kcb/periodic.htm>
Subject keywords: Materials Science, Empirical Potentials, Density Functional Theory, Energetics properties, Elastic properties
JARVIS (Joint Automated Repository for Various Integrated Simulations) is a repository designed to automate materials discovery using classical force-field, density functional theory, machine learning calculations and experiments. The Force-field section of JARVIS (JARVIS-FF) consists of thousands of automated LAMMPS based force-field calculations... show more
Harvested from CHIMAD MDF [mrr.materialsdatafa...](#)
- Elemental vacancy diffusion database from high-throughput first-principles calculations for fcc and hcp structures**
Angsten, Thomas/University of Wisconsin-Madison, Mayeshiba, Tam/University of Wisconsin-Madison, Wu, Henry/University of Wisconsin-Madison, Morgan, Dane/University of Wisconsin-Madison - NIST
<http://hdl.handle.net/11256/76>
Subject keywords: Materials Science, Bulk Diffusion, Density Functional Theory
This work demonstrates how databases of diffusion-related properties can be developed from high-throughput ab initio calculations. The formation and migration energies for vacancies of all adequately stable pure elements in both the face-centered cubic (fcc) and hexagonal close packing (hcp) crystal structures were determined using ab initio calcul... show more
Harvested from CHIMAD MDF [mrr.materialsdatafa...](#)
- Supplementary information for a study of DFT+U in the context of BiFeO3**
John Kane Shenton/University College London, London, London, United Kingdom - GitHub.com
<https://github.com/JohnKaneShenton/colin-2014-06-01-01-dft+u-in-the-context-of-bifeo3>

PDR Preservation Service

- Public Data Repository specializes in “science-ready” data products
 - Standard Reference Data (SRD), Data Publications
 - Collections of files, ancillary data (previews, figures), metadata
- Preservation Service
 - package data for long-term storage
 - Standards-based
 - Packaging based on BagIt: NIST Preservation Profile

BagIt Packaging Standard

- Basics:
 - A Bag = a structured directory hierarchy (can be serialized, compressed)
 - **data** directory preserves native data organization
 - Manifest file provides checksums for each data file
 - Packaging metadata
 - Allows customization—profiles—for storing metadata and other data
- Requirements: Transmission versus Preservation
 - Aggregates heterogeneous files, Information preserving
 - Self-describing, validate-able, corruption-detection
 - Scalability to large collections:
 - Transmission: **fetch.txt** lists URLs for externally retrieving files that are part of collection
 - Preservation: should not rely on existence of external service

Multibag BagIt Profile

- Splits a logical bag over multiple component bags
- Drivers:
 - Data products that are large: total size or number of files
 - Allows for more efficient storage and retrieval
 - Efficient revisions to preserved data products
 - Create new component bags that contain only the parts that have changed
- Profile formalizes relationships between bags
 - “Head bag”: points to other component bags
 - Defines how to recombine components to reconstruct complete bag
 - Revisions create a new “head bag”

Materials Resource Registry: Supporting Data Discovery

<https://materials.registry.nist.gov/>

Search for Resources

electrical x properties x

All Resources Organizations Data Collections Datasets Services Informational Sites Software

Search criteria used (Clear all):
Type x Material Type x 5 results

TYPE (Clear)

- Organization (0)
- Collection (0)
- Dataset (2)
- Service (0)
- Software (0)
- Web Site (1)

ORIGIN OF DATA (Clear)

MATERIAL TYPE (Clear)

- biological (0)
- biomaterials (1)
- ceramics (2)
- metals and alloys (2)
- metamaterials (0)
- molecular fluids (0)
- organic compounds (1)
- organometallics (1)
- polymers (4)
- semiconductors (2)

STRUCTURAL FEATURE (Clear)

PROPERTY ADDRESSED (Clear)

MatWeb
MatWeb
<http://www.matweb.com/index.aspx>
Subject keyword(s): engineering, material properties, database, suppliers
"The heart of MatWeb's services is our searchable online database of engineering materials. We have over 115,000 data sheets in our collection and have many powerful search tools available to help our users find the materials information that they need. While we have a variety of services that we offer to companies in the engineering community, our... show more

Microelectronics Packaging Materials Database
CINDAS LLC
<https://cindasdata.com/products/mpmd>
Subject keyword(s): electronics packaging materials, material properties
"The MPMD contains data and information on thermal, mechanical, electrical and physical properties of electronics packaging materials, and it is available in a Web-based format. The database is continually updated and expanded. The MPMD contains over 1,025 materials, 363 properties, and contains approximately 22,500 data curves."

JARVIS - Joint Automated Repository for Various Integrated Simulations
Kamal Choudhary/National Institute of Standards and Technology, Faical Y. Congo/National Institute of Standards and Technology, Chandler A. Becker, Francesca M. Tavazza/National Institute of Standards and Technology - NIST
<http://www.ctcms.nist.gov/~knc6/periodic.html>
Subject keyword(s): Materials Science, Empirical Potentials, Density Functional Theory, Energetics properties, Elastic properties
JARVIS (Joint Automated Repository for Various Integrated Simulations) is a repository designed to automate materials discovery using classical force-field, density functional theory, machine learning calculations and experiments. The Force-Field section of JARVIS (JARVIS-FF) consists of thousands of automated LAMMPS based force-field calculations... show more
↳ Harvested from ChimAd MDF mirr.materialsdatafa...

Elemental vacancy diffusion database from high-throughput first-principles calculations for fcc and hcp structures
Angsten, Thomas/University of Wisconsin-Madison, Mayeshiba, Tam/University of Wisconsin-Madison, Wu, Henry/University of Wisconsin-Madison, Morgan, Dane/University of Wisconsin-Madison - NIST
<http://hdl.handle.net/11256/76>
Subject keyword(s): Materials Science, Bulk Diffusion, Density Functional Theory
This work demonstrates how databases of diffusion-related properties can be developed from high-throughput ab initio calculations. The formation and migration energies for vacancies of all adequately stable pure elements in both the face-centered cubic (fcc) and hexagonal close packing (hcp) crystal structures were determined using ab initio calcul... show more
↳ Harvested from ChimAd MDF mirr.materialsdatafa...

Supplementary information for a study of DFT+U in the context of BiFeO3
John Kane Shenton/University College London: London, London, United Kingdom - Github.com
<https://github.com/JohnKaneShenton/colondon>

NIST Home Services Dashboard Logout Help Contact

Materials Resource Registry

Add New Resource

My Resources

Resource

Local ID: JJWTWSDURKLF1JBV4WIT

Status: Active

Identity

Resource Name: NIST Chemistry WebBook

Alternate Name:

Version:

Identifier:

Logo:

Providers

Publisher: National Institute of Standards and Technology

Publication Year:

Creator:

Contributor:

Date:

Contact

Name: NIST

Postal Address:

Email Address: webbook11@nist.gov

Phone Number:

Time Zone:

Content

Description:

Applicability to Material Science: experiments informatics and data science simulations theory

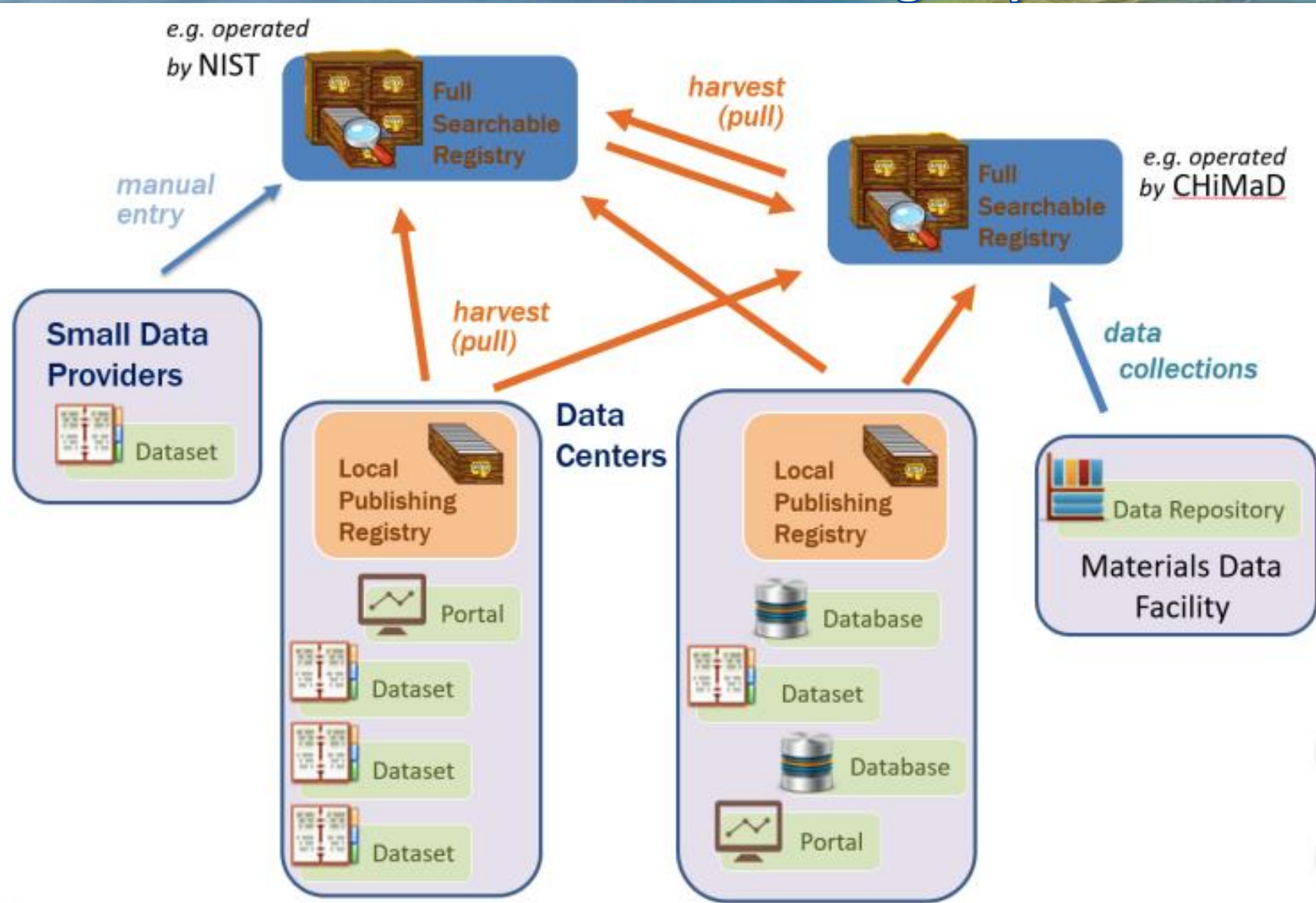
Material Type: biological biomaterials ceramics carbides cements nitrides oxides perovskites silicates metals and alloys metamaterials molecular fluids organic compounds organometallics polymers semiconductors

Structural Feature: Property Addressed: Experimental Characterization Methods:

*This site provides thermochemical, thermophysical, and ion energetics data compiled by NIST under the Standard Reference Database (SRD) program.

Materials Resource Registry:

- Reg
- C
- M
- h
- S
- se



www.nist.gov/

tions via

n

Resource Metadata

<https://materials.registry.nist.gov/>

Metadata framework supports extensibility, evolution

- Generic core metadata + extensions
- Domain-specific extensions
 - Clients can ignore extensions it doesn't understand
- **Materials Vocabulary**
 - Modest but practical: 3 tiers of detail
 - See <https://www.rd-alliance.org/materials-vocabulary-draft-21-mar-2017>
 - SKOS version available
- Based on XML, XML Schema
 - Use of XSD Types to support plug-in extensions
- JSON-based version of techniques used in the PDR
 - JSON Schema: schema definition, types, validation
 - JSON-LD: tie definitions to community semantics